Journal of Innovative Image Processing (JIIP) (2020)
Vol.02/ No. 02
Pages: 102-109
https://www.irojournals.com/iroiip/
DOI: https://doi.org/10.36548/jiip.2020.2.005

RETRIEVAL OF COMPLEX IMAGES USING VISUAL SALIENCY GUIDED COGNITIVE CLASSIFICATION

Dr. T. Vijayakumar,
Professor,
Department of ECE,
GNIT, Hyderabad, India.
Email id: vishal_16278@yahoo.co.in

Mr. R. Vinothkanna,
Department of ECE,
Vivekanandha College of Technology for Women,
Namakal, India.
Email: rvinothkannaphd@gmail.com

Abstract: Data storage via multimedia technology is more preferred as the information in multimedia contain rich meanings and are concise when compared to the traditional textual information. However, efficient information retrieval is a crucial factor in such storage. This paper presents a cognitive classification based visual saliency guided model for the efficient retrieval of information from multimedia data storage. The Itti visual saliency model is described here for generation of an overall saliency map with the integration of color saliency, intensity and direction maps. Multi-feature fusion paradigms are used for providing clear description of the image pattern. The definition is based on two stages namely complexity based on cognitive load and classification of complexity at a cognitive level. The image retrieval system is finalized by integrating a group sparse logistic regression model. In complex scenarios, the baselines are overcome by the proposed system when tested on multiple databased as compared to other state-of-the-art models.

Keywords: Feature Extraction; Image Retrieval; Complex Image; Visual Saliency;

1. Introduction

There is a massive increase in the amount of multimedia information with the rapid advancements in the internet technology and increasing popularity of digital and multimedia devices for information collection and storage [1]. Service activities and public information exchange makes use of the massive multimedia content in the form of video and images every day. Multimedia information is more user friendly and contain rich and concise information when compared to the traditional text data [2]. Hence people are more attracted towards the usage of multimedia information. There is a continuous and explosive growth in the image databases as the multimedia technology is also cheaper, more convenient and faster with new image processing and digital technologies [3]. These technologies have penetrated vastly into the lives of people due to their great convenience and vast library of digital information

ISSN: 2582- 4252 (online)

Journal of Innovative Image Processing (JIIP) (2020)
Vol.02/ No. 02
Pages: 102-109
https://www.irojournals.com/iroiip/
DOI: https://doi.org/10.36548/jiip.2020.2.005

which is easy to explore. An urgent problem in this scenario is the accuracy and speed of retrieval of images [4].

Since the early 1970s, three significant stages are crossed by the image retrieval models according to the literature review [5]. The images in the image library are appended with text description manually in the traditional systems where the desired image is then found with a keyword and the keyword matching system returns the results from the database. Ease of implementation is the major advantage of the text-based information retrieval (TBIR) approach [6]. The major shortcomings of this system include (a) the size of the image library and manual workload annotation are proportional to each other, (b) the conception of image in the mind of people may differ with the manual annotation on subjective basis [7].

Manual annotation is possible if the size of the image database is small however with the rapid increase in the volume of image databases, it is completely impossible to perform manual annotation [8]. With the revolutionary growth in the fields of machine learning, database technology and artificial intelligence, there is a quick growth in content-based image retrieval (CBIR) technology since the 1990s [9]. The traditional text description of images and keyword based matching techniques are replaced by the automatic extraction of visual features and characteristics of the images based on the similarity between measurements. The color features, texture features, space features and shape feature of the image are to be precisely extracted for accurate implementation of the CBIR system [10].

The solution to overcome the semantic gap can be classified into three major categories namely (a) Improving the accuracy of image description by improving the image annotation accuracy, (b) extraction of the most effective feature of the image for description of the image, and (c) enhancement of the overall performance by inclusion of relevant feedback to the image. A novel relevant feedback model is constructed based on the integration of the complexity level of the image and visual saliency [11]. The preliminary query result is returned by the system on entering the query request. The result is marked by the user with features like setting of irrelevant or relevant mark. Based on the signature of the user, the system learns the information and then the features of the image are modified and the result is returned. These procedures are repeated until the desired results are obtained. The weight adjustment technique and query vector modification are the categories of classification of the related feedback. The eigenvector weight adjustment, and modification as well as reflection of query request of the user is made possible with the query vector using these techniques [12].

The major factor involved in implementation of this approach is tagging of images which can be performed using two significant procedures namely machine tagging and artificial tagging. Better description of the image content and high accuracy are the advantages of artificial tagging and manual annotation [13]. However, the process is laborious and time consuming. Selection of labels and keywords is a cumbersome process. This process does not undergo any corrections, resulting in occurrence of wrong marks due to human errors. Tagging the images with machine learning algorithms for semantic labelling using computers is made possible using machine automatic labelling and machine

ISSN: 2582- 4252 (online)

Journal of Innovative Image Processing (JIIP) (2020)
Vol.02/ No. 02
Pages: 102-109
https://www.irojournals.com/iroiip/
DOI: https://doi.org/10.36548/jiip.2020.2.005

tagging techniques. Thorough analysis of the image is performed in the fundamental image and the vision characteristic vector is generated and extracted. The analysis of the characteristic image is performed using machine learning techniques and parameterization evaluation technique. Finally, the new image is tagged with the trained model [14].

2. Prerequisites

2.1 Saliency Detection

Visual attention is driven basically by two common techniques namely top-down and bottom-up. An image attracts the attention of people and emerges automatically if the surrounding area of the image has sufficient contrast in the visual scene area with bottom-up visual attention driven by the image data. This feature is termed as prominence. Current visual goals, expectations and brain knowledge are some of the high-level cognitive factors that drive the visual attention in top-down mode. The general saliency map generation and visual attention model is presented in Figure 1. The visual prominence is a phenomenon where the attention of the viewer is automatically attracted by certain features in the image. Further, more attention is required to focus on other objects or locations in the visual scene. The final gaze is a parallel combination of both these mechanisms.

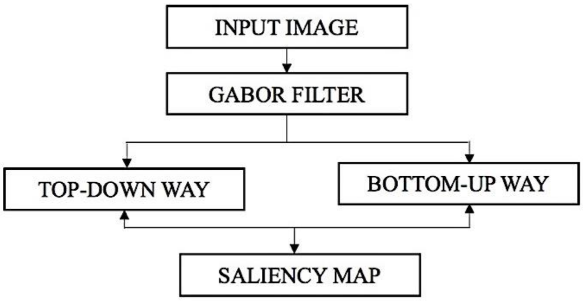

Figure 1: Framework for detection of visual saliency

Here, the Itti model is introduced in which the color and intensity saliency map as well as the direction is integrated for generation of an overall saliency map. The lowest base image is produced by the convolution input image using the Gabor kernel by the Gabor pyramid. The Gabor kernel is used for convolution of the i^{th} image to the layer of Gaussian pyramid. The kernel is chosen from $\theta \in \{0,45,90,135\}$ that are the four major directions which emphasizes the directionality of the Gabor core. A direction Gabor pyramid can be produced by processing an image of high intensity by the directions of the Gabor kernel. Four Gabor pyramids are obtained by using the Gabor nucleus's four nucleus. Equation 1 represents the feature maps.

ISSN: 2582- 4252 (online)

Journal of Innovative Image Processing (JIIP) (2020)
Vol.02/ No. 02
Pages: 102-109
https://www.irojournals.com/iroiip/
DOI: https://doi.org/10.36548/jiip.2020.2.005

$$O(k,s,\theta)=|O(k,\theta)\Theta O(s,\theta)| \text{ ------ } (1)$$

The map direction is obtained by the Gabor pyramid image represented by O(k,s,θ) based on the operation. A 24 directions feature map can be obtained by the four directions of the pyramid.

3. Image Retrieval Model

The images are matched under some geometric constraints by considering the invariant gray point differences and retrieving the image on the basis of the point of interest. The retrieval effect is determined by the point of interest using the key attribute of shape feature. The deviation is located by a large point that exists on the point of detection due to the diverse interests that makes it challenging for accurate characterization and completion of the image characteristics and shape. This reduces the efficiency of the image retrieval algorithm. It is not ideal to use the salient point of the image by image retrieval algorithm and hence other means of image representation is essential than with the significant point. The accurate representation of the image shape is challenging and the image contains the special point which is also the significant point.

The shape of the image is characterized in the traditional techniques by means of the significant point. For retrieval of images, the shape representation is prominently irregular. For the purpose of image retrieval, application of significant points that are reproducible and require detection are applied. This feature enables repeated detection of the image under varied conditions like change in visual angle or brightness in the same position. Accurate location of the significant points is challenging due to their sensitivity to noise and illumination. The edge of the object is focused mainly for the significant points. Along with the local information of the image, the image of the image as well as the global information that assists in image retrieval.

4. Results and Discussion

Experimental verification is conducted using the Corel image library which is discussed in this section. From the Corel image library, thousand images are downloaded and an image database is created. The size of the images is either 256×384 or 384×256. The images are further categorized into 10 classes of 100 images each. The possibility of retrieval of images can be increased by means of the salient region by generating the natural image saliency map. A testing database with some sample images is used for this purpose. Based on this analysis, it is found that the proposed method offers satisfactory results in terms of accuracy of image retrieval. Figure 2 and Figure 3 presents the performance precision in terms of average normal and the recall performance respectively. The experimental results are constructed by obtaining the average values from a set of 50 test images for each model. Testing of various methods in terms of precision averages at standard recall is performed. Comparison of the proposed model with

ISSN: 2582- 4252 (online)

Journal of Innovative Image Processing (JIIP) (2020)
Vol.02/ No. 02
Pages: 102-109
https://www.irojournals.com/iroiip/
DOI: https://doi.org/10.36548/jiip.2020.2.005

traditional and other state-of-the-art models is performed. Figure 4 provides the different levels of complexity of images and the corresponding levels of precision averages at standard recall.

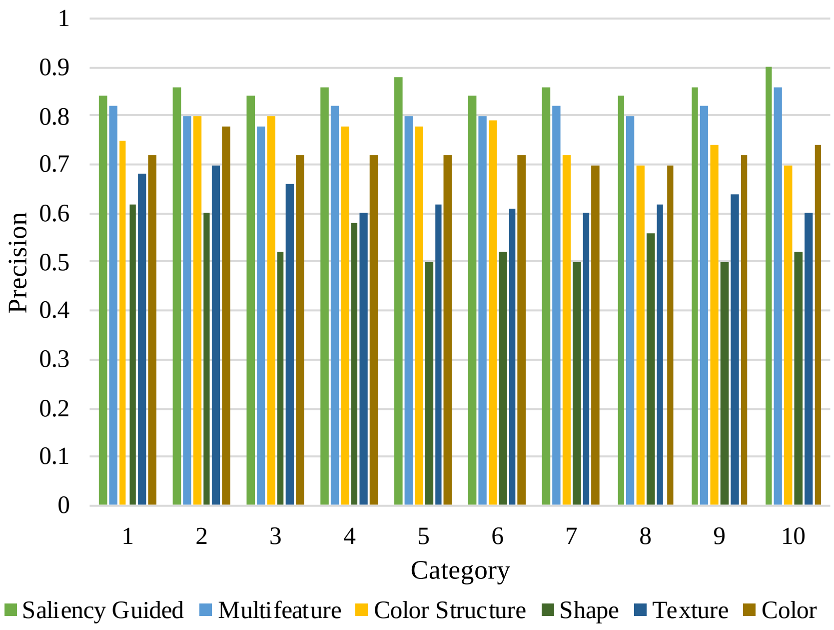

Figure 2: Average normal precision performance

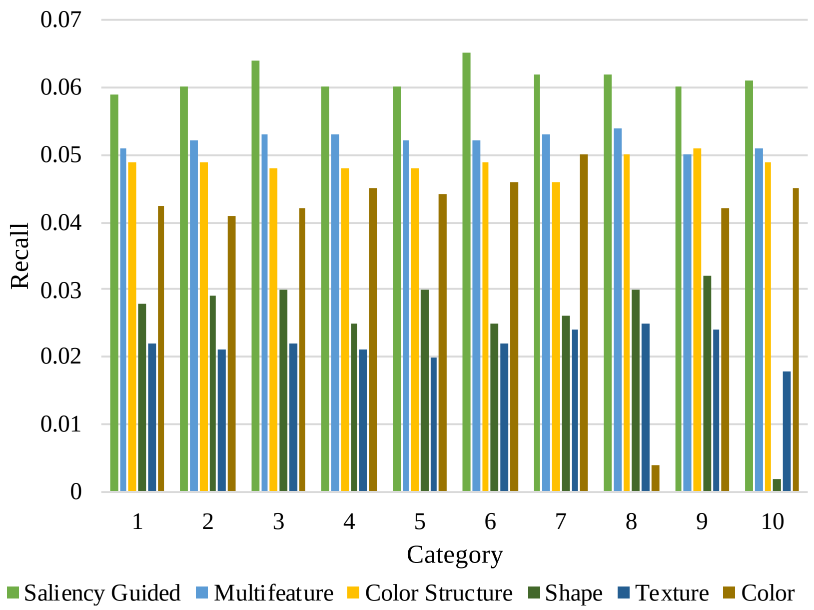

Figure 3: Average normal recall performance

ISSN: 2582- 4252 (online)

Journal of Innovative Image Processing (JIIP) (2020)
Vol.02/ No. 02
Pages: 102-109
https://www.irojournals.com/iroiip/
DOI: https://doi.org/10.36548/jiip.2020.2.005

The average precision values are compared for various methods for images of similar or varied levels of complexity for the purpose of quantitative analysis. Varied techniques with varied levels of image dimensions are used for analyzing the average precision values for considering the image dimensions. Based on these results, it is observed that, on comparison of the proposed algorithm with other existing systems, improved performance is obtained. Accuracy of results is also high on implementation of the proposed algorithm based on the analysis results across various categories of images. The performance of average normal recall and normal precision has reached 0.06 and 85% respectively. The proposed offers an average precision value of 0.8. With varied levels of complexity, the average precision value ranges at [0.60, 0.80] and under varied levels of image dimensions, the average value of precision range at [0.67, 0.81].

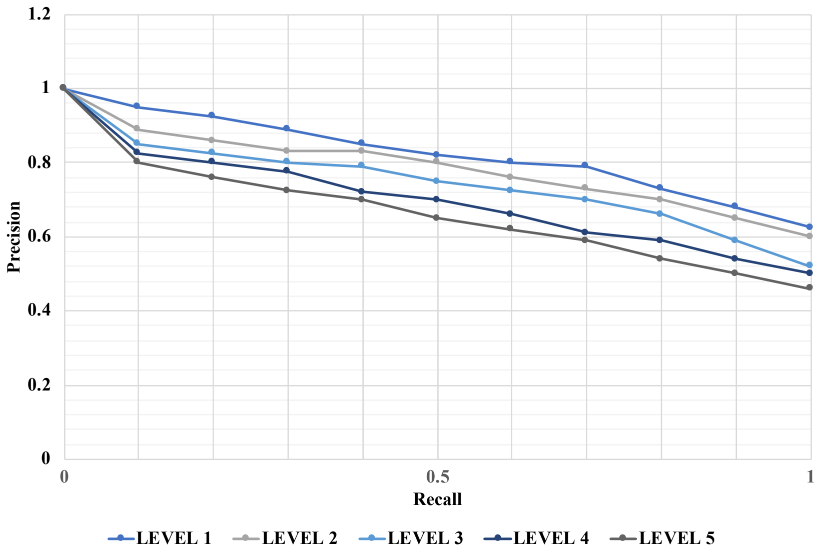

Figure 4: Precision averages at standard recall levels for different levels of image complexity

5. Conclusion

This paper presents a novel image retrieval model using visual saliency guided complex based on cognitive classification. The solution to overcome the semantic gap can be classified into three major categories namely (a) Improving the accuracy of image description by improving the image annotation accuracy, (b) extraction of the most effective feature of the image for description of the image, and (c) enhancement of the overall performance by inclusion of relevant feedback to the image. A novel relevant feedback model is constructed based on the integration of the complexity level of the image and visual saliency. The consistency among the search result and the query image is verified in geometric terms by making use of any two-point Euclidean distance in the significant region between the matching features. The effectiveness of the presented model is validated by the experimental results. Experimental verification of the image is conducted using the Corel image library in the simulation.

ISSN: 2582- 4252 (online)

Journal of Innovative Image Processing (JIIP) (2020)
Vol.02/ No. 02
Pages: 102-109
https://www.irojournals.com/iroiip/
DOI: https://doi.org/10.36548/jiip.2020.2.005

From the Corel image library, thousand images are downloaded and an image database is created. The size of the images is either 256×384 or 384×256. The images are further categorized into 10 classes of 100 images each. On testing and comparison of the average recall performance and normal precision, it is found that there is an increase in the rates of the proposed model. Future work focuses developing a robust model that verifies the similarity in the image properties on a multiple database platform.

References:

[1] Yoshida, M., & Veale, R. (2014). Saliency-guided neural prosthesis for visual attention: Design and simulation. Neuroscience research, 78, 90-94.

[2] Erdem, E., & Erdem, A. (2013). Visual saliency estimation by nonlinearly integrating features using region covariances. Journal of vision, 13(4), 11-11.

[3] Wang, Q., Zheng, W., & Piramuthu, R. (2016). Grab: Visual saliency via novel graph model and background priors. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 535-543).

[4] Calvo, M. G., & Nummenmaa, L. (2011). Time course of discrimination between emotional facial expressions: The role of visual saliency. Vision research, 51(15), 1751-1759.

[5] Jayashree, S. and D. A. Janeera. "Real-Time Fire Detection, Alerting and Suppression System using Live Video Surveillance." (2016).

[6] Liang, X., Zhang, J., Tian, Q., Li, J., & Zhuo, L. (2018, April). A Saliency Guided Shallow Convolutional Neural Network for Traffic Signs Retrieval. In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 340-345). IEEE.

[7] Danko, A. S. (2015). Visual saliency estimation A pre-attentive cognitive and context-aware approach. State University of New York at Albany.

[8] Wei, S., Liao, L., Li, J., Zheng, Q., Yang, F., & Zhao, Y. (2019). Saliency Inside: Learning Attentive CNNs for Content-Based Image Retrieval. IEEE Transactions on Image Processing, 28(9), 4580-4593.

[9] Ruth Anita Shirley D, Ranjani K, Gokulalakshmi Arunachalam, Janeera D.A., "Distributed Gardening System Using Object Recognition and Visual Servoing" In International Conference on Inventive Communication and Computational Technologies [ICICCT 2020], Springer, India, 2020.

[10] Muratov, O. (2013). Visual saliency detection and its application to image retrieval (Doctoral dissertation, University of Trento).

[11] Evangelopoulos, G., Zlatintsi, A., Potamianos, A., Maragos, P., Rapantzikos, K., Skoumas, G., & Avrithis, Y. (2013). Multimodal saliency and fusion for movie summarization based on aural, visual, and textual attention. IEEE Transactions on Multimedia, 15(7), 1553-1568.

[12] Andrushia, A. D., & Thangarjan, R. (2018). Saliency-based image compression using walsh–hadamard transform (WHT). In Biologically rationalized computing techniques for image processing applications (pp. 21-42). Springer, Cham.

[13] Chandy, A. (2019). RGBD Analysis for Finding the Different Stages of Maturity of Fruits in Farming. Journal of Innovative Image Processing (JIIP), 1(02), 111-121.

ISSN: 2582- 4252 (online)

Journal of Innovative Image Processing (JIIP) (2020)
Vol.02/ No. 02
Pages: 102-109
https://www.irojournals.com/iroiip/
DOI: https://doi.org/10.36548/jiip.2020.2.005

[14] Manoharan, S. (2020). Performance Analysis of Clustering Based Image Segmentation Techniques. Journal of Innovative Image Processing (JIIP), 2(01), 14-24.

Authors Biography

Dr. T. Vijayakumar, is currently working as Professor, in Department of ECE, at GNIT, Hyderabad, India. His research are includes Computer Vision, Motion Analysis, Stereo Vision, Object Recognition, computer graphics, photo interpretation, image retrieval, Embedded Image Processing and Real-time image and video processing applications.

Mr. R. Vinothkanna, is presently working as the professor in Department of ECE, at Vivekanandha College of Technology for Women, Namakal, India. His major area of research are Imaging Technology, Pattern Recognition, Biomedical Imaging, Biometrics , Health care Applications ,Capsule Networks

ISSN: 2582- 4252 (online)